42390.P10853 *Patent*

UNITED STATES PATENT APPLICATION

FOR

Omnivergent Stereo Image Capture

INVENTORS:

Fernando Martins

Prepared by

Steven D. Yates
Reg. No. 42,242
(503) 264-6589

Express Mail mailing label number: EL546136697US

Omnivergent Stereo Image Capture

Field of the Invention

The invention generally relates to image capture, and more particularly to capturing omnivergent stereo image pairs for compactly representing 360° panoramic images of an environment surrounding a capture device.

Background

There are many known techniques for generating three-dimensional images of a surrounding environment. Methods referred to as "360 panorama", "Panoramic stereo imaging" and "omnivergent imaging" concern distinct techniques for creating such three-dimensional images. Some techniques utilize one or more image capturing devices, e.g., cameras or other input source, to define stereo images corresponding to an environment about the image capturing devices.

For example, the conference paper *Stereo panorama with a single camera*, Peleg and Ben-Ezra, in Proc. Computer Vision and Pattern Recognition Conf., pp. 395-401 (1999), discusses creating mosaic images from a rotating camera, dealing with incident problems of parallax and scale changes, and using a single camera to create two stereo panoramic views, one for each eye, through multiple viewpoint image projections.

The paper *Stereo reconstruction from multiperspective panoramas*, Heung-Yeung Shum and Richard Szeliski, in IEEE Int'l Conference on Computer Vision, pp. 14-21 vol. 1 (1999), discusses computing depth maps from a collection of images,

where camera motion is constrained to planar concentric circles. The resulting collection of regular perspective images is sampled into a set of multiple perspective panoramas, and depth maps can be computed from the sampled images.

The conference paper Omnivergent Stereo by Heung-Yeung Shum, and Seitz, in IEEE Int'l Conference on Computer Vision, pp. 22-29 vol. 1 (1999), discusses a virtual sensor for 3D image construction, where instead of using planar perspective images that collect many rays at a fixed viewpoint, omnivergent cameras are instead used to collect a small number of selected rays at many different viewpoints. The collected 2D manifold of rays is arranged into two multiple-perspective images allowing for stereo reconstruction, and, according to the authors, 3D models can be reconstructed with minimal error since every point in the model is defined with respect to two rays having maximum vergence angle. However, this document fails to teach how a physical capture device can be created to implement the techniques discussed in the document.

It will be appreciated by those skilled in the art that these three references are presented for exemplary purposes to illustrate current state of the art, and to show lack of knowledge in the art as to how to effectively build an omnivergent stereo image capture device not suffering from limitations discussed in these references.

Brief Description Of The Drawings

The features and advantages of the present invention will become apparent from the following detailed description of the present invention in which:

FIG. 1 illustrates a top view according to one embodiment of an image recorder having a cylindrical region defined by the rotation path of a deflector.

FIG. 2 illustrates a rotation of the FIG. 1 deflector.

FIG. 3 illustrates an alternate embodiment utilizing four deflectors and four sensors.

FIG. 4 illustrates a perspective view of the FIG. 1 embodiment.

FIG. 5 illustrates a suitable computing environment in which certain aspects of the invention may be implemented.

Detailed Description

The following detailed description assumes reader familiarity with the mathematics and principles of stereo and omnivergent imaging.

FIG. 1 illustrates a top view according to one embodiment of an omnivergent stereo recorder. A cylindrical region **100** is defined by the rotation path of a deflector **102** about an axis of rotation **104** that is perpendicular to a rotation plane **106**. In one embodiment, on, adjacent or proximate to the axis of rotation is a receptor **108** comprising an array of image sensors that is also perpendicular to the rotation plane. In illustrated embodiments, the receptor comprises two one-dimensional columns **110**, **112** of sensors. However, it will be appreciated that other arrangements may be used.

As illustrated, there are two inputs **114**, **116** to the deflector **102** that are tangential to the cylindrical region **100**. In one embodiment, the inputs are directly opposing each other so as to form a 180° angle between them. In one embodiment, the deflector is a prism and the inputs are light rays. It will be appreciated, however, other visible and/or non-visible electromagnetic energy may be converged with the deflector **102**, and the deflector will have properties appropriate to the electromagnetic energy

being reflected. For example, the deflector may be a physical deflector - in the case of a prism deflecting light. For certain other forms of electromagnetic energy, the deflector may be a generated deflection field, such as a magnetic field.

When the two inputs **114**, **116** reach the deflector **102**, they are deflected so as to converge on the sensors **110**, **112**. Assuming a left-to-right arrangement in the figure, as illustrated, the left input **114** is deflected towards the left sensor **110**, and the right input **116** is deflected towards the right sensor **112**. For each rotational position of the deflector-sensor assembly, the sensors are used to record the received input for the given rotational position. After recording the input, as illustrated in FIG. 2, the sensors can be rotated **200** to a new rotational position and inputs collected again. This process of rotation, collection of input, and continuing rotation can be repeated continuously.

The rate of rotation can be accelerated or decreased according to temporal sampling needs. For example, to generate a live broadcast of three-dimensional data the rotation speed needs to be at least 1800 rotations per minute (RPM) to maintain a frame rate of 30 frames per second (FPS). It will be appreciated that a 30 FPS can also be achieved with a lower rotational speed by having multiple deflectors **102** and receptors **108**. For example, as illustrated in FIG. 3, four deflectors **300** can be arranged with four sensors **302** to allow four concurrent samples to be taken for a given rotation position. This embodiment reduces rotation speed to 450 RPM for maintaining a frame rate of 30 FPS, since a full revolution of samples is now collected in a quarter turn. Alternatively, this embodiment allows quadrupling the frame rate. It will be appreciated that an arbitrary number of sensors and deflectors may be used.

To create an omnivergent stereo image pair, a first image is defined by compiling all leftward inputs, i.e. the set of inputs **114,** captured over a complete revolution of a deflector-sensor assembly (e.g., a combination of a deflector **102** and a sensor **108**) A second image is similarly defined using the rightward inputs, i.e. the set of inputs **116** captured over a complete revolution of the deflector-sensor assembly. These two images form an omnivergent stereo image pair which compactly represents the three dimensional structure of the environment surrounding the invention.

FIG. 4 illustrates a perspective view of the FIG. 1 embodiment. As in FIG. 1, illustrated are the cylindrical region **100** defined by the rotation path of the deflector **102** about the axis of rotation **104**. As can now be seen in FIG. 4, the deflector and sensors are mechanically related. As the deflector is rotated, the array of sensors **400**, **402** (corresponding to FIG. 1 items **110**, **112**) rotates in tandem with the deflector. During rotation, the sensors constantly receive the deflected input **114**, **116**.

It will be appreciated that sensor arrays **400**, **402**, can be arbitrarily dense and have diverse physical arrangements depending on the technology used to create the sensor arrays. The more sensors there are, the better the vertical resolution of captured images, and when combined with arbitrary rotation speeds, captured image data can be arbitrarily detailed and dense. It will also be appreciated that the distance between the deflector **102** and the axis of rotation **104** may be arbitrarily adjusted so as to determine a desired depth of field and maximum vergence angle for imaging a particular environment.

Once the omnivergent stereo images have been captured and determined, one can then select an arbitrary viewpoint from within the cylindrical region **100** and use image-based rendering techniques to render a virtual image with respect to the arbitrarily selected viewpoint. Selection and rendering may be performed in real time, allowing for significant advances in telepresence, video conferencing applications, three dimensional gaming applications, and other applications in which it would be useful to allow users to select arbitrary viewpoints in a scene. The invention enables users to experience holographic television without the need for holograms as intermediate storage. The limitation is that the viewpoints can only be selected from within the cylinder **100.**

In one embodiment, multiple omnivergent stereo image capturing devices (not illustrated) are used to determine omnivergent stereo images from multiple viewpoints. These multiple omnivergent stereo images are then synthesized into a large viewpoint selection region comprising the individual cylindrical regions of the multiple capturing devices. In one embodiment, the multiple capturing devices are arranged so that their cylindrical regions abut each other. In another embodiment, synthesizing includes computing values for intervening spaces between cylindrical regions.

FIG. 5 and the following discussion are intended to provide a brief, general description of a suitable computing environment in which certain aspects of the illustrated invention may be implemented. It will be appreciated by one skilled in the art that the invention may be used in applications such as simulated holographic television videoconferencing, three-dimensional gaming, virtual reality environment, to capture

and/or react to movement within an environment, e.g., user gesticulation, or other application desiring three-dimensional representations of an environment.

Towards these ends, the invention may incorporate or be incorporated within, controlled by, or otherwise communicatively coupled with a machine **500** having system bus **502** for coupling various machine components. Typically, attached to the bus are one or more processors **504**, a memory **506** (e.g., RAM, ROM), storage devices **508**, a video interface **510**, and input/output interface ports **512**. The system may include embedded controllers, such as programmable logic devices or gate arrays, Application Specific Integrated Circuits (ASIC), single-chip computers, etc.

The system may operate in a networked environment using physical and/or logical connections to one or more remote systems **514**, **516** through a network interface **518**, modem **520**, or other pathway. Systems may be interconnected by way of a wired or wireless network **522**, including an intranet, the Internet, local area networks, wide area networks, cellular, cable, laser, satellite, microwave, "Blue Tooth" type networks, optical, infrared, or other carrier.

The invention may be described by reference to program modules or instructions for performing tasks or implementing abstract data types, e.g., procedures, functions, data structures, application programs, etc., that may be stored in memory **506** and/or storage devices **508** and associated volatile and non-volatile storage media, e.g., magnetic, optical, biological, or other storage, as well as transmission environments such as network **522** over which program modules may be delivered in the form of packets, serial data, parallel data, or other transmission format. Program modules or

instructions may be stored in machine accessible media, including wired and wirelessly accessible media.

Thus, for example, assuming a three-dimensional videoconferencing or multi-player game, machine **500** and remote devices **514**, **516** operate to determine an omnivergent stereo image of their respective environments. These stereo images can then be shared among users of the machines **500**, **514**, **516** to allow respective users to selectively define viewpoints within other users' environments. It will be appreciated that remote machines **514**, **516** may be configured like machine **500**, and therefore include many or all of the elements discussed for machine. It should also be appreciated that machines **500**, **514**, **516** may be embodied within a single device, or separate communicatively-coupled components.

Having described and illustrated the principles of the invention with reference to illustrated embodiments, it will be recognized that the illustrated embodiments can be modified in arrangement and detail without departing from such principles. And, even though the foregoing discussion has focused on particular embodiments, it is understood other configurations are contemplated. In particular, even though expressions such as "in one embodiment," "in another embodiment," or the like are used herein, these phrases are meant to generally reference embodiment possibilities, and are not intended to limit the invention to particular embodiment configurations. As used herein, these terms may reference the same or different embodiments, and unless indicated otherwise, embodiments are combinable into other embodiments.

Consequently, in view of the wide variety of permutations to the above-described embodiments, the detailed description is intended to be illustrative only, and should not

be taken as limiting the scope of the invention. What is claimed as the invention, therefore, is all such modifications as may come within the scope and spirit of the following claims and equivalents thereto.